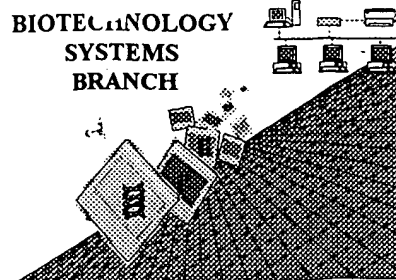# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/936,883_

Source: _PCT/09_

Date Processed by STIC: _10/4/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

RAW SEQUENCE LISTING   DATE: 10/04/2001
PATENT APPLICATION: US/09/936,883   TIME: 10:11:25

Does Not Comply
Corrected Diskette Needed

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\10042001\I936883.raw

pp 1-4

```
 4 <110> APPLICANT: Miyata, Toshio
 6 <120> TITLE OF INVENTION: A METHOD FOR DETECTING MEGSIN PROTEIN
 7       AND USE THEREOF
 9 <130> FILE REFERENCE: SHIM012
11 <140> CURRENT APPLICATION NUMBER: US/09/936,883
12 <141> CURRENT FILING DATE: 2001-09-07
14 <150> PRIOR APPLICATION NUMBER: PCT/JP00/01646
15 <151> PRIOR FILING DATE: 2000-03-17
17 <150> PRIOR APPLICATION NUMBER: 11/75305
18 <151> PRIOR FILING DATE: 1999-03-19
20 <150> PRIOR APPLICATION NUMBER: 11/306623
21 <151> PRIOR FILING DATE: 1999-10-28
23 <160> NUMBER OF SEQ ID NOS: 21
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
```

ERRORED SEQUENCES

```
     27 <210> SEQ ID NO: 1
     28 <211> LENGTH: 1867
     29 <212> TYPE: DNA
     30 <213> ORGANISM: Homo sapiens
     32 <400> SEQUENCE: 1
     33 atggcctccc ttgctgcagc aaatgcagag ttttgcttca acctgttcmt aasruaaaaa 60
E--> 34 aasnaaguhc yshasnuhag agagatggat gacaatcaag gaaatggaaa tgtgttcttt 120
E--> 35 tcctctargg umtasasasn gngyasngya snvahhsrsr ctgagcctct tcgctgccct 180
E--> 36 ggccctggtc cgcttgggcg ctcaagatus ruhaaaauaa uvaargugya agnasgactc 240
E--> 37 cctctctcag attgataagt tgcttcatgt taacactgcc tcaassrusr gnasysuuhs 300
E--> 38 vaasnthraa srggatatgg aaactcttct aatagtcagt cagggctcca gtctcaactg 360
E--> 39 gytyrgyasn srsrasnsrg nsrgyugnsr gnuaaaagag tttttttctga tataaatgca 420
E--> 40 tcccacaagg attatgatct cysargvahs rasasnaasr hsysastyra suagcattgt 480
E--> 41 gaatgggctt tttgctgaaa aagtgtatgg ctttcataag srvaasngyu haaguysvat 540
     42 yrgyhhsysg actacattga gtgtgccgaa aaattatacg atgccaaagt ggagcgaast 600
     43 yrgucysaag uysutyrasa aysvaguarg gttgacttta cgaatcattt agaagacact 660
E--> 44 agacgtaata ttaataagva ashthrasnh suguasthra rgargasnas nystgggttg 720
E--> 45 aaaatgaaac acatggcaaa atcaagaacg tgattggtga atrvaguasn guthrhsgyy 780
E--> 46 sysasnvagy guggtggcat aagctcatct gctgtaatgg tgctggtgaa tgctgtgtac 840
E--> 47 gygysrsrsr aavamtvauv aasnaavaty rttcaaaggc aagtggcaat cagccttcac 900
E--> 48 caagagcgaa accataaath ysgyystrgn sraahthrys srguthrasn tgccatttca 960
     49 aatctcccaa gtgctctggg aaggcagtcg ccatgatgcy shshyssrry scyssrgyys 1020
     50 aavaaamtmt catcaggaac ggaagttcaa tttgtctgtt attgaggacc catcaatghs 1080
E--> 51 gnguargysh asnusrvagu asrsrmtaag attcttgagc tcagatacaa tggtggcata 1140
E--> 52 aacatgtacg ttctgysugu uargtyrasn gygyasnmtt yrvauctgcc tgagaatgac 1200
E--> 53 ctctctgaaa ttgaaaacaa actgaccttt cagurguasn asusrgugua snysuthrhg 1260
E--> 54 naatctaatg gaatggacca atccaaggcg aatgacctct aagtatgtta snumtgutrt 1320
E--> 55 hrasnrarga rgmtthrsry styrvagagg tatttttttcc tcagttcaag atagagaaga 1380
E--> 56 attatgaaat gaaaguvahh rgnhysguys asntyrgumt yscaatattt gagagcccta 1440
```

) see item 9 on Error Summary Sheet

RAW SEQUENCE LISTING   DATE: 10/04/2001
PATENT APPLICATION: US/09/936,883 TIME: 10:11:25

    Input Set : **A:\seqlist.txt**
    Output Set: **N:\CRF3\10042001\I936883.raw**

```
E--> 57 gggctgaaag atatctttga tgaatccaaa gntyruarga augyuysash asgusrysgc 1500      ) item 9
     58 agatctctct gggattgctt cgggggggtcg tctgtatata tcaaggaaas usrgyaasrg 1560
     59 ygyargutyr srargatgat gcacaaatct tacatagagg tcactgagga gggcaccgag 1620
     60 gctmtmthsy ssrtyrguva thrgugugyt hrguaaactg ctgccacagg aagtaatatt 1680      ) item 9
E--> 61 gtagaaaagc aactccctca gtccthraaa athrgysras nvaguysgnu rgnsracgct 1740
     62 gtttagagct gaccacccat tcctatttgt tatcaggaag gatthruhar gaaashsrhu 1800
     63 hvaargysas gacatcatct tattcagtgg caaagtttct tgcccttgaa suhsrgyysv 1860
     64 asrcysr                                                             1867
    313 <210> SEQ ID NO: 18
    314 <211> LENGTH: 1938
    315 <212> TYPE: RNA
    316 <213> ORGANISM: Rattus norvegicus
    318 <400> SEQUENCE: 18
E--> 319 tttcaaaatg gcctcccttg ctgcagcaaa tgcagattt ggcttcgacm taasruaaaa 60
E--> 320 aaasmaaguh gyhasttatt cagagagatg gatagtagtc aaggaaacgg aaatgtattc 120
E--> 321 ttcunarggu mtassrsrgn gyasngyasn vahhtcttcc ctgagcatct tcactgccct 180
E--> 322 gagcctaatc cgtttggggtg ctsrsrusrh thraausrua rgugyaacga ggtgactgtn 240
E--> 323 nncgtcagat tgacaaggcc ctgcacttta tctccarggy ascysaaarg gnasysaauh 300
E--> 324 shsrccatca agacaaggga attcatcgaa cagtcagcta ggactgcaat atrsrarggn 360
E--> 325 gyasnsrsra snsrgnugyu gntyrcaatt gaaaagagtt cttgctgaca taaactcatc 420
E--> 326 tcataaggat nnngnuysar gvauaaasas nsrsrhsysa saaaaactca gcattgccaa 480
E--> 327 tggagttttt gcagagaaag tatttgattt tysusraaas ngyvahaagu ysvahashca 540
E--> 328 taagagctat atggagtgtg ctgaaaactt atacaatgct aaagtghsys srtyrmtguc 600
E--> 329 ysaaguasnu tyrasnaays vagaaagagt tgattttaca aatgatatac aagaaaccag 660
E--> 330 atttaaaatt guargvaash thrasnasgn guthrarghy saataaatgg attgaaaatg 720
E--> 331 aaacacatgg caaaatcaag aaggtgttga snystrguas nguthrhsgy ysysysvaug 780
E--> 332 gggacagcag cctcagctca tcagctgtca tggtgctagt gaatgctgya ssrsrusrsr 840
E--> 333 sraavamtva uvaasnaagt ttacttcaaa ggcaagtgga aatcggcctt caccaagagt 900
E--> 334 gataccvaty rhysgyystr yssraahthr yssrasthrc tcagttgcca tttcaggtct 960
E--> 335 cccagcggtc ctggaaaagc agttaatusr cyshshargs rrsrgyrgyy saavaasnat 1020
E--> 336 gatgcatcaa gaacggaggt tcaatttgtc taccattcag gagccamtmt hsgnguarga 1080
E--> 337 rghasnusrt hrgngurcca atgcagattc ttgagctaca atatcatggt ggcataagca 1140
E--> 338 tgtacrmtgn uguugntyrh sgygysrmtt yratcatgtt gcccgaggat gacctatccg 1200
E--> 339 aaattgaaag caagctgagt mturguasas usrgugusry susrttccag aatctaatgg 1260
E--> 340 actggacaaa tagcaggaag atgaaatctc aghgnasnum tastrthras nsrargysmt 1320
E--> 341 yssrgntatg tgaatgtgtt tctcccccag ttcaagatag agaaagatta tgaatyrvaa 1380
E--> 342 snvahurgnh ysguysasty rguatgagga gccacttgaa atctgtaggc ttggaagaca 1440
E--> 343 tctttgttga gmtargsrhs uyssrvagyu guashvagut ccagggctga tctgtctgga 1500
E--> 344 attgcctctg gaggtcgtct ctatgtasra rgaaasusrg yaasrgygya rgutyrvatc 1560
E--> 345 aaagctaatg cacaagtccc tcatagaggt ctcagaagaa ggcaccsrys umthsyssru 1620
E--> 346 guvasrgugu gythrgaggc aactgctgcc acagaaagta acatcgttga aaagctactt 1680
E--> 347 cctguaathr aaaathrgus rasnvaguys uurgaatcca cggtgttcag agctgaccgc 1740
E--> 348 cccttctgt ttgtcattag ggusrthrva hargaaasar grhuhvaarg aagaatggca 1800
E--> 349 tcatcttatt tactggcaaa gtctcgtgtc ctysasngyu hthrgyysva srcysrtgaa 1860
E--> 350 attctatttg gtttttccata cactaacagg catgaagaaa catcataagt gaatagaatt 1920
E--> 351 gtaattggaa gtacatgg                                                 1938
    353 <210> SEQ ID NO: 19
    354 <211> LENGTH: 380
    355 <212> TYPE: PRT
```

(handwritten annotations)

see item 9

→ "t's not allowed in an RNA sequence

Per 1.823 of Sequence Rules, a combined DNA/RNA sequence requires a <212 response of DNA; also, explain in <2207-<2237 section that sequence is a combined DNA/RNA 10/4/01

RAW SEQUENCE LISTING   DATE: 10/04/2001
PATENT APPLICATION: US/09/936,883   TIME: 10:11:25

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\10042001\I936883.raw

```
356 <213> ORGANISM: Rattus norvegicus
358 <400> SEQUENCE: 19
359 Met Ala Ser Leu Ala Ala Ala Asn Ala Glu Phe Gly Phe Asp Leu Phe
360  1               5               10              15
361 Arg Glu Met Asp Ser Ser Gln Gly Asn Gly Asn Val Phe Phe Ser Ser
362          20              25              30
363 Leu Ser Ile Phe Thr Ala Leu Ser Leu Ile Arg Leu Gly Ala Arg Gly
364          35              40              45
```
E-->
```
365 Asp Cys Xaa Arg Gln Ile Asp Lys Ala Leu His Phe Ile Ser Pro Ser
366          50              55              60
367 Arg Gln Gly Asn Ser Ser Asn Ser Gln Leu Gly Leu Gln Tyr Gln Leu
368 65              70              75              80
```
E-->
```
369 Lys Arg Val Leu Ala Asp Ile Asn Ser Ser His Lys Asp Xaa Lys Leu
370              85              90              95
371 Ser Ile Ala Asn Gly Val Phe Ala Glu Lys Val Phe Asp Phe His Lys
372              100             105             110
373 Ser Tyr Met Glu Cys Ala Glu Asn Leu Tyr Asn Ala Lys Val Glu Arg
374          115             120             125
375 Val Asp Phe Thr Asn Asp Ile Gln Glu Thr Arg Phe Lys Ile Asn Lys
376          130             135             140
377 Trp Ile Glu Asn Glu Thr His Gly Lys Ile Lys Lys Val Leu Gly Asp
378 145             150             155             160
379 Ser Ser Leu Ser Ser Ser Ala Val Met Val Leu Val Asn Ala Val Tyr
380              165             170             175
381 Phe Lys Gly Lys Trp Lys Ser Ala Phe Thr Lys Ser Asp Thr Leu Ser
382              180             185             190
383 Cys His Phe Arg Ser Pro Ser Gly Pro Gly Lys Ala Val Asn Met Met
384          195             200             205
385 His Gln Glu Arg Arg Phe Asn Leu Ser Thr Ile Gln Glu Pro Pro Met
386          210             215             220
387 Gln Ile Leu Glu Leu Gln Tyr His Gly Gly Ile Ser Met Tyr Ile Met
388 225             230             235             240
389 Leu Pro Glu Asp Asp Leu Ser Glu Ile Glu Ser Lys Leu Ser Phe Gln
390              245             250             255
391 Asn Leu Met Asp Trp Thr Asn Ser Arg Lys Met Lys Ser Gln Tyr Val
392              260             265             270
393 Asn Val Phe Leu Pro Gln Phe Lys Ile Glu Lys Asp Tyr Glu Met Arg
394          275             280             285
395 Ser His Leu Lys Ser Val Gly Leu Glu Asp Ile Phe Val Glu Ser Arg
396          290             295             300
397 Ala Asp Leu Ser Gly Ile Ala Ser Gly Gly Arg Leu Tyr Val Ser Lys
398 305             310             315             320
399 Leu Met His Lys Ser Leu Ile Glu Val Ser Glu Glu Gly Thr Glu Ala
400          325             330             335
401 Thr Ala Ala Thr Glu Ser Asn Ile Val Glu Lys Leu Leu Pro Glu Ser
402          340             345             350
403 Thr Val Phe Arg Ala Asp Arg Pro Phe Leu Phe Val Ile Arg Lys Asn
404          355             360             365
405 Gly Ile Ile Leu Phe Thr Gly Lys Val Ser Cys Pro
```

*see item 9
on Error
Summary
Sheet*

```
        406      370                375                380
        409 <210> SEQ ID NO: 20
        410 <211> LENGTH: 1848
        411 <212> TYPE: DNA
        412 <213> ORGANISM: Mus Musculus
        414 <400> SEQUENCE: 20
        415 ttcgacttat tcagagagat ggatagtagc caaggaaatg gaaatgtaha suhargqumt 60
E-->    416 assrsrgngy asngyasnva ttcttctctt ccctgagcat cttcactgcc ctgaccctaa 120
        417 tccgtctghh srsrusrhth raauthruar gugqtgctcg aggtgactgt gcacgtcaga 180
E-->    418 ttgacaaggc actgcacttt gyaaarggya scysaaargg nasysaauhs haacatacca 240
E-->    419 tcaagacaag gaaactcatc taataatcag ccaggactta snrsrarggn gyasnsrsra 300
E-->    420 snasngnrgy ucagtatcaa ttgaaaagag ttcttgctga cataaactca tctcataagg 360
E-->    421 ntyrgnuysa rgvauaaasa snsrsrhsys gattatgaac tcagcattgc cactggagtt 420
        422 tttgcagaaa aagtctatas tyrguusraa thrgyvahaa guysvatyrg actttcataa 480
E-->    423 gaactacatt gagtgtgctg aaaacttata caatgctash hsysasntyr gucysaagua 540
E-->    424 snutyrasna aaaagtggaa agagttgact tcacaaatga tgtacaagat accagattty 600
E-->    425 svaguargva ashthrasna svagnasthr arghaaaatt aataaatgga ttgaaaatga 660
E-->    426 gacacatgga aagatcaaga agysasnyst rguasnguth rhsgyysysy sgtgttgggg 720
        427 gacagcagcc tcagctcgtc ggctgtcatg gtgctggtgv augyassrsr usrsrsraav 780
E-->    428 amtvauvaaa cgctgtttac ttcaaaggca aatggaaatc ggccttcacc aagactasna 840
        429 avatyrhysg yystryssra ahthrysthr gataccctca gttgccgttt taggtctccc 900
        430 acgtgtcctg gaaaagtaas thrusrcysa rghargsrrt hrcysrgyys vagttaatat 960
E-->    431 gatgcatcaa gaacggeggt tcaatttgtc taccattcag vaasnmtmth sgnguargar 1020
E-->    432 ghasnusrth rgncagccac caatgcaggt tcttgagctc caatatcatg gtggcataag 1080
E-->    433 cgnrrmtgnv auguugntyr hsgygysrat gtacattatg ctgcctgagg atggcctatg 1140
        434 tgaaattgaa agcaagmtty rmturguasg yucysgugus rysctgagtt tccagaatct 1200
E-->    435 gatggactgg accaatagga ggaaaatgaa ausrhgnasn umtastrthr asnargargy 1260
E-->    436 smtystctca gtatgtgaac gtgtttctcc cccagttcaa gatagagaag aatsrgntyr 1320
E-->    437 vaasnvahur gnhysguysa sntatgaaat gacgcaccac ttgaaatcct taggcttgaa 1380
        438 agatatcttt tyrgumtthr hshsuyssru gyuysashga tgagtccagt gcagatctct 1440
        439 ctggaattgc ctctggaggt cgtctcasgu srsraaasus rgyaasrgyg yargutacgt 1500
        440 atcaaagcta atgcacaagt cattcataga ggtctcagag gagtyrvasr ysumthsyss 1560
        441 rhquvasrgu guggcactga agccactgct gccacagaaa ataacattgt tgaaaagcag 1620
E-->    442 gythrguaat hraaaathrg uasnasnvag uysgnettee tgagtccaca gtgttcagag 1680
        443 ccgaccgccc ctttctgttt gtcurqusrt hrvahargaa asargrhuhv aatcaagaag 1740
E-->    444 aatgacatca tcttattttac tggcaaagtc tcttgtccty sysasnasuh thrgyysvas 1800
        445 rcysrtgaaa ttcgatttgg tttcctatac agtaacaggc atcaagaa              1848
```

Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding
explanation is presented in the <220> to <223> fields of
each sequence using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/936,883**

DATE: 10/04/2001
TIME: 10:11:26

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\10042001\I936883.raw**

```
L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:34 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:1
M:340 Repeated in SeqNo=1
L:132 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:3
L:132 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:3
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:137 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:144 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:4
L:144 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:4
L:144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:156 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:5
L:156 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:5
L:156 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:161 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:218 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:319 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:15
L:320 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:18
L:320 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:320 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:27
M:340 Repeated in SeqNo=18
L:321 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:321 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:39
L:322 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:11
L:322 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:50
L:323 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:9
L:323 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:59
L:324 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:9
L:324 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:68
L:325 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:11
L:325 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:79
L:326 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:6
L:326 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:85
L:327 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:14
L:327 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:99
L:328 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:15
L:328 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:114
L:329 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:10
L:329 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:124
L:330 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:330 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:136
L:331 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:7
L:331 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:143
L:332 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:11
L:332 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:154
L:333 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:11
L:333 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:165
L:334 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:13
```

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/936,883**

DATE: 10/04/2001
TIME: 10:11:26

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\10042001\I936883.raw**

```
L:334 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:178
L:335 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:6
L:335 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:184
L:336 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:13
L:336 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:197
L:337 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:337 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:209
L:338 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:13
L:338 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:222
L:339 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:10
L:339 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:232
L:340 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:9
L:340 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:241
L:341 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:16
L:341 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:257
L:342 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:9
L:342 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:266
L:343 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:13
L:343 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:279
L:344 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:344 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:291
L:345 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:8
L:345 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:299
L:346 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:10
L:346 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:309
L:347 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:8
L:347 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:317
L:348 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:12
L:348 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:329
L:349 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:15
L:349 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:344
L:350 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:19
L:350 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:363
L:351 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:5
L:351 M:321 E: (1) "t" not allowed in RNA Sequence, NUMBER OF INVALID 't' KEYS:368
L:365 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:19
M:340 Repeated in SeqNo=19
L:416 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:20
M:340 Repeated in SeqNo=20
```

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: *09/936,883*

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAF

1____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3____Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5____Variable Length

Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7____Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8____Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9____Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11____Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12____PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001